

54



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
 United States Patent and Trademark Office
 Address: COMMISSIONER FOR PATENTS
 P.O. Box 1450
 Alexandria, Virginia 22313-1450
 www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/058,104	01/29/2002	Keiichi Chihara	B0104T	1202
32628	7590	03/17/2005	EXAMINER	
HAUPTMAN KANESAKA BERNER PATENT AGENTS SUITE 300, 1700 DIAGONAL RD ALEXANDRIA, VA 22314-2848			ALBERTALLI, BRIAN LOUIS	
			ART UNIT	PAPER NUMBER
			2655	

DATE MAILED: 03/17/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/058,104	CHIHARA, KEIICHI	
	Examiner	Art Unit	
	Brian L Albertalli	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-14 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-14 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 29 January 2002 is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. ____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. ____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| Paper No(s)/Mail Date ____ | 6) <input type="checkbox"/> Other: ____ |

DETAILED ACTION

Drawings

1. The drawings are objected to because:
 - a) In Fig. 2, "condour" near label 802 should be --contour--.
 - b) In Fig. 7, "tablel" in element 1408 should be --table--.
 - c) In Fig. 8, "condour" in step 114 should be --contour--.

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

Claim Objections

2. Claim 3 is objected to because of the following informalities:

Lines 12-14 of the claim state that the pitch contour determination unit determines a pitch contour using either the *duration rule table* or the *duration prediction table*. However, it appears from the specification that this is a typographical error. The specification discloses that the pitch contour determination unit (Fig. 2) uses prediction table 909 and rule table 910, which appear to be separate from the duration prediction and rule tables (Fig. 3, 1006 and 1007). Furthermore, it is not clear how the pitch contour could be generated from tables built for generating a phoneme duration.

Therefore, for the purposes of examination, the duration rule table and duration prediction table, as claimed in claim 3, have been interpreted herein as the separate prediction table (909) and rule table (910), as disclosed in the specification.

Appropriate correction is required.

Claim Rejections - 35 USC § 112

3. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

4. Claim 8 recites the limitation "said threshold" in lines 1-2 of the claim. There is insufficient antecedent basis for this limitation in the claim.

For the purposes of examination, "said threshold" has been interpreted herein as the point at which the determination step for the switch of parent claim 7 is made.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 1 and 2 are rejected under 35 U.S.C. 103(a) as being unpatentable over applicant's admitted prior art, in view of Otsuka (U.S. Patent 6,546,367).

The applicant's admitted prior art discloses a method of controlling high-speed reading in a text-to-speech conversion system including a text analysis module for generating a phoneme and prosody character string from an input text (Fig. 15, 101); a prosody generation module for generating a synthesis parameter of at least a voice

segment, a phoneme duration, and a fundamental frequency for said phoneme and prosody character string (Fig. 16); a voice segment dictionary in which voice segments as a source of voice are registered (Fig. 15, 105); and a speech generation module for generating a synthetic waveform by waveform superimposition by referring to said voice segment dictionary (Fig. 15, 103).

Applicant's admitted prior art does not disclose:

the step of providing said prosody generation module with a phoneme duration determination unit that includes both a duration rule table containing empirically found phoneme durations and a duration prediction table containing phoneme durations predicted by statistical analysis and determines a phoneme duration by using, when a user-designated utterance speed exceeds a maximum utterance speed threshold, said duration rule table and, when said threshold is not exceeded, said duration prediction table.

Otsuka discloses a method comprising the step of:

providing said prosody generation module with a phoneme duration determination unit (Fig. 2, phoneme duration setting unit 5) that includes both a duration rule table containing empirically found phoneme durations (Fig. 4, threshold values θ) and a duration prediction table containing phoneme durations predicted by statistical analysis (Fig. 4, average value μ , standard deviation value σ , and minimum value d) and determines a phoneme duration by using, when a user-designated utterance speed exceeds a threshold, said duration rule table and, when said threshold is not exceeded, said duration prediction table.

See Figure 5. In step 107, an initial phoneme production time is determined dependent on the total speech production time T (thereby determining an initial rate of speech, column 3, line 63 to column 4, line 2 and column 4, lines 15-17). If this initial phoneme production time is less than the empirically found phoneme durations (threshold values θ), the threshold values are used as the phoneme duration (column 6, lines 8-10). Otherwise, the durations predicted by statistical analysis are used (average value μ , standard deviation value σ , and minimum value d are used to set a phoneme duration with the most probable value, column 7, lines 22-27). The threshold values used are necessarily the maximum utterance speed, because any initial phoneme duration that is less than the threshold duration will be set to the threshold duration (producing speech at the minimum phoneme duration is equivalent to producing speech at the maximum utterance speed).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the applicant's admitted prior art to use, when a user-designated utterance speed exceeds a threshold, said duration rule table and, when said threshold is not exceeded, said duration prediction table, in order to realize a natural phoneme duration regardless of the speech production time (utterance speed), as taught by Otsuka (column 14, lines 30-34).

7. Claims 3 and 4 are rejected under 35 U.S.C. 103(a) as being unpatentable over applicant's admitted prior art, in view of Otsuka, and further in view of Vermeulen et al. (U.S. Patent 6,810,379).

As discussed in reference to claim 1, above, the applicant's admitted prior art discloses all of the features of the instant claims, except:

the step of providing said prosody generation module with a pitch contour determination unit that has both an empirically found rule table and a prediction table predicted by statistical analysis and determines a pitch contour by determining both accent and phrase components with, when a user-designated utterance speed exceeds a maximum utterance speed threshold, said pitch contour rule table and, when said threshold is not exceeded, said pitch contour prediction table.

Otsuka discloses a method of switching between a statistical table and a rule-based table depending on the selected utterance speed (Fig. 5).

Neither the applicant's admitted prior art nor Otsuka disclose using those tables to determine a pitch contour.

Vermeulen et al. disclose that text-to-speech systems can use both rule based and statistical models (column 2, lines 10-11).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of the applicant's admitted prior art and Otsuka to include a statistical table and a rule table for the pitch contour, and to use the rule table when a maximum utterance speed threshold had been exceeded, in order to realize a natural pitch contour regardless of the speech production time (utterance speed).

8. Claims 5 and 6 are rejected under 35 U.S.C. 103(a) as being unpatentable over applicant's admitted prior art, in view of Hara et al. (U.S. Patent 5,615,300).

As discussed in reference to claim 1, above, the applicant's admitted prior art discloses all of the features of the instant claims, except:

the step of providing said prosody generation module with a sound quality coefficient determination unit that has a sound quality conversion coefficient table for changing said voice segment to switch sound quality and selects from said sound quality conversion coefficient table such a coefficient that sound quality does not change when a user-designated utterance speed exceeds a maximum utterance speed threshold.

Hara et al. disclose a method comprising:

the step of providing said prosody generation module with a sound quality coefficient determination unit (Fig. 3, mode selector 21) that has a sound quality conversion coefficient table for changing said voice segment to switch sound quality (mode selector 21 selects the order of the cepstral parameters, which determines the sound quality, column 11, line 60 to column 12, line 3 and column 10, lines 31-33) and selects from said sound quality conversion coefficient table such a coefficient that sound quality does not change when a CPU activity ration exceeds a threshold (if the CPU activity ratio is greater than 30%, the order N is set to 6).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the applicant's admitted prior art to not change the quality when the utterance speed exceeded a maximum utterance speed threshold, since the utterance

Art Unit: 2655

speed would be directly related to the CPU activity (increases in speed would increase CPU activity), and this would ensure that the speech quality selection would not be set so low that the output speech became unintelligible.

9. Claims 7-9 are rejected under 35 U.S.C. 103(a) as being unpatentable over applicant's admitted prior art, in view of Rye (*Speech Synthesis at Higher Speaking Rates*).

In regard to claims 7 and 8, as discussed in reference to claim 1, above, the applicant's admitted prior art discloses all of the features of the instant claims, except:

the step of providing said prosody generation module with both a pitch contour correction unit for outputting a pitch contour corrected according to an intonation level designated by the user and a switch for determining whether a base pitch is added to said pitch contour corrected according to said user-designated utterance speed.

Rye discloses that the selection of overall voice pitch affects intelligibility at high speaking rates (page 3, High Voiced Pitch or Low Pitch section).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify applicant's admitted prior art to include a pitch contour correction unit for determining whether to add a base pitch to the pitch contour in order to prevent a loss in intelligibility.

In regard to claim 9, the applicant's admitted prior art discloses said pitch contour correction unit performs a pitch contour generation process that includes a phrase

component calculation process in which all phrases of an input sentence are processed by calculating a phrase component by statistical analysis, and a process in which all words in said input sentence are processed by calculating an accent component by statistical analysis and correcting said accent component according to said user-designated intonation level (user designated intonation determines phrase components A_{pi} and accent component A_{aj} , which are used by base pitch addition unit 505 according to the statistical analysis of Equation (1), Fig. 18 and page 6, line 25 to page 7, line 12).

The applicant's admitted prior art does not disclose using a user-designated utterance speed to determine whether to correct the phrase and accent components with the contour correction unit, or making the phrase and accent components zero.

Rye discloses that the selection of overall voice pitch affects intelligibility at high speaking rates and user's choice for voice pitch will affect intelligibility at high speeds (page 3, High Voiced Pitch or Low Pitch section and page 4, Discussion section, lines 6-7).

It would have been obvious to one of ordinary skill in the art at the time of invention to make the phrase and accent components zero when a user-designated utterance speed was exceeded, so that the user's designated intonation level would not adversely affect the intelligibility of the output speech at high speaking rates.

10. Claims 10 and 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over applicant's admitted prior art.

As discussed in reference to claim 1, above, the applicant's admitted prior art discloses all of the features of the instant claims, except:

the step of providing said speech generation module with signal sound generation means for inserting a signal sound between sentences to indicate an end of a sentence when a user-designated utterance speed exceeds a maximum utterance speed threshold.

Official notice is taken that it is notoriously well known and recognized in the art that indexing spoken speech with signal sounds (audible tones) helps a listener to easily understand where there are important transitions in the text being spoken (as in tone indexing).

It would have been obvious to one of ordinary skill in the art at the time of invention to insert a signal sound between sentences when the utterance speed exceeded a maximum utterance speed threshold, so the user would easily understand the transition between sentences.

11. Claims 12-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over applicant's admitted prior art, in view of Walsh (U.S. Patent Application Publication 2003/0014253).

In regard to claims 12 and 13, as discussed in reference to claim 1, above, the applicant's admitted prior art discloses all of the features of the instant claims, except:

the step of providing said prosody generation module with a phoneme duration determination unit for performing a process in which when a user-designated utterance

speed exceeds a maximum utterance speed threshold, an utterance speed of at least a leading word in a sentence is returned to a normal utterance speed.

Walsh discloses a method of variably changing the phoneme duration of words in a text-to-speech system (Fig. 7A and 7B and paragraphs 48-50) and further acknowledges that greater emphasis should be given between sentences (page 2, paragraph 30, lines 24-29).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the applicant's admitted prior art to return the leading word of a sentence to normal utterance speed when the user-designated utterance speed exceeded a maximum utterance speed threshold, in order to reduce the playing duration without reducing the comprehensibility of the message, as taught by Walsh (page 5, paragraph 50).

In regard to claim 14, the applicant's admitted prior art does not disclose correcting the length of a vowel or vowels of the word.

As discussed above in reference to claims 12 and 13, Walsh discloses not correcting a phoneme duration (leaving the duration at normal speed) and correcting a phoneme duration according to the user-designated utterance speed (see Fig. 6A and 6B). Furthermore, Walsh discloses correcting the length of each phoneme, which necessarily includes the vowels of each word (page 4, paragraph 45 and Table II).

It would have been obvious to one of ordinary skill in the art at the time of invention to not correct the phoneme duration of words at the beginning of a sentence

(leave the phonemes at normal speed) and correct the phoneme duration of words when the phonemes were not at the beginning of a sentence, in order to reduce the playing duration without reducing the comprehensibility of the message, as taught by Walsh (page 5, paragraph 50).

Conclusion

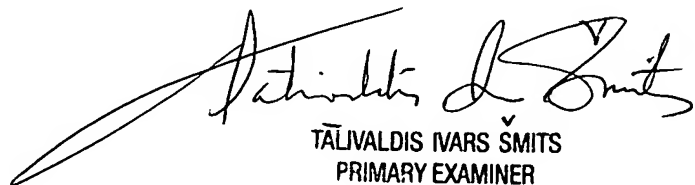
12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Yegnanarayana et al. (*Voice Simulation : Factors Affecting Quality and Naturalness*) disclose pitch contour affects speech quality. Hirschberg et al. (*Building Study Skills for Students with Vision Loss*) disclose the use of aural markers help people locate information in spoken audio. Holm et al. (U.S. Patent 6,260,016) disclose a system that uses pre-constructed prosody templates. Eide et al. (U.S. Patent 6,101,470) disclose a method for producing speech contours. Karaali et al. (U.S. Patent 5,913,194) disclose a method of using a HMM network that includes rule based transitions to create speech from text. Huang et al. (U.S. Patent 5,905,972) disclose a method of using base pitch templates. Masuzawa et al. (U.S. Patent 4,279,030) disclose a system that produces an audible tone before a speech announcement is made. Masuzawa et al. (U.S. Patent 4,700,393), Silverman (U.S. Patent 5,749,071), Vigler (U.S. Patent 5,826,231), Nishiguchi (U.S. Patent 5,926,788), and Itoh et al. (U.S. Patent 6,205,427) disclose various methods for adjusting the speaking rate of speech.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817, until March 28, 2005. After March 28, 2005, the examiner can be reached at (571) 272-7616. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 3/15/05



TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER